


LIBRARY
OF THE
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY



Digitized by the Internet Archive
in 2011 with funding from
Boston Library Consortium Member Libraries

Optimum Consumption and Portfolio Rules
in a Continuous-time Model

by

Robert C. Merton

Number 58 August 1970

This paper was supported by the National Science Foundation (grant number GS 1812). The views expressed in this paper are the author's sole responsibility, and do not reflect those of the National Science Foundation, the Department of Economics, nor of the Massachusetts Institute of Technology.

Robert C. Merton

M.I.T. March, 1970

1. Introduction. A common hypothesis about the behavior of (limited liability) asset prices in perfect markets is the random walk of returns or (in its continuous-time form) the "geometric Brownian motion" hypothesis which implies that asset prices are stationary and log-normally distributed. A number of investigators of the behavior of stock and commodity prices have questioned the accuracy of the hypothesis.⁽¹⁾ In particular, Cootner [2] and others have criticized the independent increments assumption, and Osborne [2] has examined the assumption of stationariness. Mandelbroit [2] and Fama [2] argue that stock and commodity price changes follow a stable-Paretian distribution with infinite second-moments. The non-academic literature on the stock market is also filled with theories of stock price patterns and trading rules to "beat the market", rules often called "technical analysis" or "charting", and that presupposes a departure from random price changes.

* I would like to thank P.A. Samuelson, R.M. Solow, P.A. Diamond, J.A. Mirrlees, J.S. Flemming, and D.T. Scheffman for their helpful discussions. Of course, all errors are mine. Aid from the National Science Foundation is gratefully acknowledged

(1)

For a number of interesting papers on the subject, see Cootner [2]. An excellent survey article is "Efficient Capital Markets: A Review of Theory and Empirical Work" (a draft), E.F. Fama, November, 1969.

In an earlier paper [12], I examined the continuous-time consumption-portfolio problem for an individual whose income is generated by capital gains on investments in assets with prices assumed to satisfy the "geometric Brownian motion" hypothesis. I.e. I studied $\text{Max } E \int_0^T U(C,t)dt$ where U is the instantaneous utility function; C is consumption; E is the expectation operator. Under the additional assumption of a constant relative or constant absolute risk-aversion utility function, explicit solutions for the optimal consumption and portfolio rules were derived. The changes in these optimal rules with respect to shifts in various parameters such as expected return, interest rates, and risk were examined by the technique of comparative statics.

The present paper extends these results for more general utility functions, price behavior assumptions, and for income generated also from non-capital gains sources. It is shown that if the "geometric Brownian motion" hypothesis is accepted, then a general "Separation" or "mutual fund" theorem can be proved such that, in this model, the classical Tobin mean-variance rules hold without the objectionable assumptions of quadratic utility or of normality of distributions for prices. Hence, when asset prices are generated by a geometric Brownian motion, one can work with the two-asset case without loss of generality. If the further assumption is made that the utility function of the individual is a member of the family of utility functions called the "HARA" family, explicit solutions for the optimal consumption and portfolio rules are derived and a number of theorems proved. In the last parts of the paper, the effects on the consumption and portfolio rules of alternative asset price dynamics, in which changes are neither stationary nor independent, are examined along with the effects of introducing wage income, uncertainty of life expectancy, and the possibility of default on (formerly) "risk-free" assets.

2. A digression on Itô Processes. To apply the dynamic programming technique in a continuous-time model, the state variable dynamics must be expressible as Markov stochastic processes defined over time intervals of length h , no matter how small h is. Such processes are referred to as infinitely divisible in time. The two processes of this type ⁽²⁾ are: functions of Gauss-Wiener Brownian motions which are continuous in the "space" variables and functions of Poisson processes which are discrete in the space variables. Because neither of these processes is differentiable in the usual sense, a more general type of differential equation must be developed to express the dynamics of such processes. A particular class of continuous-time Markov processes of the first type called Itô Processes are defined as the solution to the stochastic differential equation ⁽³⁾

$$(1) \quad dP = f(P,t)dt + g(P,t)dz$$

where P , f , and g are n -vectors and $z(t)$ is a n -vector of

⁽²⁾ I ignore those infinitely divisible processes with infinite moments which include those members of the stable Paretian family other than the normal.

⁽³⁾ Itô Processes are a special case of a more general class of stochastic processes called Strong diffusion processes (see Kushner [9], p.22).

(1) is a short-hand expression for the stochastic integral

$$P(t) = P(0) + \int_0^t f(P,s)ds + \int_0^t g(P,s)dz$$

where $P(t)$ is the solution to (1) with probability one.

A rigorous discussion of the meaning of a solution to equations like (1) is not presented here. Only those theorems needed for formal manipulation and solution of stochastic differential equations are in the text and these without proof. For a complete discussion of Itô Processes, see the seminal paper of Itô [7], Itô and McKean [8], and McKean [11]. For a short description and some proofs, see Kushner [9], p.12-18. For a heuristic discussion of continuous-time Markov processes in general, see Cox and Miller [3], chapter 5.

standard normal random variables. Then $dz(t)$ is called a multi-dimensional Wiener process (or Brownian motion). ⁽⁴⁾

The fundamental tool for formal manipulation and solution of stochastic processes of the Itô type is Itô's Lemma stated as follows ⁽⁵⁾

Lemma: Let $F(P_1, \dots, P_n, t)$ be a C^2 function defined on $R^n \times [0, \infty)$ and take the stochastic integrals

$$P_i(t) = P_i(0) + \int_0^t f_i(P, s) ds + \int_0^t g_i(P, s) dz_i,$$

$$i = 1, \dots, n;$$

then the time-dependent random variable $Y \equiv F$ is a stochastic integral and its stochastic differential is

$$dY = \sum_i^n \frac{\partial F}{\partial P_i} dP_i + \frac{\partial F}{\partial t} dt + \frac{1}{2} \sum_i^n \sum_j^n \frac{\partial^2 F}{\partial P_i \partial P_j} dP_i dP_j$$

where the product of the differentials $dP_i dP_j$ are defined by the multiplication rule

$$dz_i dz_j = \rho_{ij} dt \quad i, j = 1, \dots, n$$

$$dz_i dt = 0 \quad i = 1, \dots, n$$

where ρ_{ij} is the instantaneous correlation coefficient between the Wiener processes dz_i and dz_j . ⁽⁶⁾

⁽⁴⁾ dz is often referred to in the literature as "Gaussian White Noise." There are some regularity conditions imposed on the functions f and g . It is assumed throughout the paper that such conditions are satisfied. For the details, see [9] or [11].

⁽⁵⁾ See McKean [11] p.32-35 and p.44 for proofs of the Lemma in one- and n-dimensions.

⁽⁶⁾ This multiplication rule has given rise to the formalism of writing the Wiener process differentials as $dz_i = \mathcal{J}_i \sqrt{dt}$ where the \mathcal{J}_i are standard normal variates (e.g. see [3]).

Armed with Itô's Lemma, we are now able to formally differentiate most smooth functions of Brownian motions (and hence integrate stochastic differential equations of the Itô type).⁽⁷⁾

Before proceeding to the discussion of asset price behavior, another concept useful for working with Itô Processes is the differential generator (or weak infinitesimal operator) of the stochastic process $P(t)$. Define the function $G(P,t)$ by

$$(2) \quad \overset{0}{G}(P,t) \equiv \lim_{h \rightarrow 0} E_t \left[\frac{G(P(t+h), t+h) - G(P(t), t)}{h} \right]$$

when the limit exists and where " E_t " is the conditional expectation operator, conditional on knowing $P(t)$. If the $P_i(t)$ are generated by Itô Processes, then the differential generator of P , \mathcal{L}_P , is defined by

$$(3) \quad \mathcal{L}_P \equiv \sum_1^n f_i \frac{\partial}{\partial P_i} + \frac{\partial}{\partial t} + \frac{1}{2} \sum_1^n \sum_1^n a_{ij} \frac{\partial^2}{\partial P_i \partial P_j}$$

where $f = (f_1, \dots, f_n)$, $g = (g_1, \dots, g_n)$, and $a_{ij} \equiv g_i g_j \rho_{ij}$. Further, it can be shown that

$$(4) \quad \overset{0}{G}(P,t) = \mathcal{L}_P[G(P,t)]$$

G can be interpreted as the "average" or expected time rate of change of the function $G(P,t)$ and as such is the natural generalization of the ordinary time derivative for deterministic functions.⁽⁸⁾

⁽⁷⁾ Warning: derivatives (and integrals) of functions of Brownian motions are similar to, but different from, the rules for deterministic differentials and integrals. For example, if $P(t) = P(0)e^{\int_0^t dz - \frac{1}{2}t} = P(0)e^{Z(t) - Z(0) - \frac{1}{2}t}$, then $dP = Pdz$. Hence $\int_0^t \frac{dP}{P} = \int_0^t dz \neq \log(P(t)/P(0))$.

Stratonovich [15] has developed a symmetric definition of stochastic differential equations which formally follows the ordinary rules of differentiation and integration. However, this alternative to the Itô formalism will not be discussed here.

⁽⁸⁾ A heuristic method for finding the differential generator is to take the conditional expectation of dG (found by Itô's Lemma) and "divide" by dt . The result of this operation will be $\mathcal{L}_P[G]$, i.e. formally, $\frac{1}{dt} E_t(dG) = \overset{0}{G} = \mathcal{L}_P[G]$.

The " \mathcal{L}_P " operator is often called a Dynkin operator and is written as " D_P ".

3. Asset price dynamics and the budget equation.

Throughout the paper, it is assumed that all assets are of the limited liability type; that there exist continuously-trading, perfect markets with no transactions costs for all assets; that the prices per share, $\{P_i(t)\}$, are generated by Itô Processes, i.e.

$$(5) \quad \frac{dP_i}{P_i} = \alpha_i(P, t)dt + \sigma_i(P, t)dz_i$$

where we define $\alpha_i \equiv f_i/P_i$ and $\sigma_i \equiv g_i/P_i$ and α_i is the instantaneous conditional expected percentage change in price per unit time and σ_i^2 is the instantaneous conditional variance per unit time. In the particular case where the "geometric Brownian motion hypothesis is assumed to hold for asset prices, α_i and σ_i will be constants. For this case, prices will be stationarily and log-normally distributed and it will be shown that this assumption about asset prices simplifies the continuous-time model in the same way that the assumption of normality of prices simplifies the static one-period portfolio model.

To derive the correct budget equation, it is necessary to examine the discrete-time formulation of the model and then to take limits carefully to obtain the continuous-time form. Consider a period model with periods of length h where all income is generated by capital gains and wealth, $W(t)$ and $P_i(t)$ are known at the beginning of period t . Let the decision variables be indexed such that the indices coincide with the period in which the decisions are implemented. Namely, let

$N_i(t) \equiv$ number of shares of asset i purchased during
period t , i.e. between t and $t + h$

(6) and

$C(t) \equiv$ amount of consumption per unit time during
period t

The model assumes that the individual "comes into" period t with wealth invested in assets so that

$$(7) \quad W(t) = \sum_1^n N_i(t-h)P_i(t)$$

Notice that it is $N_i(t-h)$ because $N_i(t-h)$ is the number of shares purchased for the portfolio in period $(t-h)$ and it is $P_i(t)$ because $P_i(t)$ is the current value of a share of the i^{th} asset. The amount of consumption for the period, $C(t)h$, and the new portfolio, $N_i(t)$, are simultaneously chosen, and if it is assumed that all trades are made at (known) current prices, then we have that

$$(8) \quad -C(t)h = \sum_1^n [N_i(t) - N_i(t-h)]P_i(t)$$

The "dice" are rolled and a new set of prices is determined, $P_i(t+h)$, and the value of portfolio is now $\sum_1^n N_i(t)P_i(t+h)$. So the individual "comes into" period $(t+h)$ with wealth $W(t+h) = \sum_1^n N_i(t)P_i(t+h)$ and the process continues.

Incrementing (7) and (8) by h to eliminate backward differences, we have that

$$\begin{aligned} (9) \quad -C(t+h)h &= \sum_1^n [N_i(t+h) - N_i(t)]P_i(t+h) \\ &= \sum_1^n [N_i(t+h) - N_i(t)][P_i(t+h) - P_i(t)] \\ &\quad + \sum_1^n [N_i(t+h) - N_i(t)]P_i(t) \end{aligned}$$

and

$$(10) \quad W(t+h) = \sum_1^n N_i(t)P_i(t+h)$$

Taking the limits as $h \rightarrow 0$, ⁽⁹⁾ we arrive at the continuous version of (9) and (10),

$$(9') \quad -C(t)dt = \sum_1^n dN_i(t) dP_i(t) + \sum_1^n dN_i(t) P_i(t)$$

and

$$(10') \quad W(t) = \sum_1^n N_i(t)P_i(t)$$

⁽⁹⁾ We use here the result that Itô Processes are right-continuous (see [9], p.15) and hence $P_i(t)$ and $W(t)$ are right-continuous. It is assumed that $C(t)$ is a right-continuous function and throughout the paper, the choice of $C(t)$ is restricted to this class of functions.

Using Itô's Lemma, we differentiate (10') to get

$$(11) \quad dW = \sum_1^n N_i dP_i + \sum_1^n dN_i P_i + \sum_1^n dN_i dP_i$$

The last two terms, $\sum_1^n dN_i P_i + \sum_1^n dN_i dP_i$, are the net value of additions to wealth from sources other than capital gains.⁽¹⁰⁾ Hence, if $dy(t) =$ (possibly stochastic) instantaneous flow of non-capital gains (wage) income, then we have that

$$(12) \quad dy - C(t)dt = \sum_1^n dN_i P_i + \sum_1^n dN_i dP_i$$

From (11) and (12), the budget or accumulation equation is written as

$$(13) \quad dW = \sum_1^n N_i(t) dP_i + dy - C(t)dt$$

It is advantageous to eliminate $N_i(t)$ from (13) by defining a new variable, $w_i(t) \equiv N_i(t)P_i(t)/W(t)$, the percentage of wealth invested in the i^{th} asset at time t . Substituting for dP_i/P_i from (5), we can write (13) as

$$(14) \quad dW = \sum_1^n w_i W \alpha_i dt - C dt + dy + \sum_1^n w_i W \sigma_i dz_i$$

where, by definition, $\sum_1^n w_i \equiv 1$.⁽¹¹⁾

Until section seven, it will be assumed that $dy \equiv 0$, i.e. all income is derived from capital gains on assets. If one of the n -assets is "risk-free" (by convention, the n^{th} asset), then $\sigma_n = 0$, the instantaneous rate of return, α_n , will be called r , and (14) is re-written as

$$(14') \quad dW = \sum_1^m w_i (\alpha_i - r) W dt + (rW - C)dt + dy + \sum_1^m w_i \sigma_i dz_i$$

⁽¹⁰⁾ This result follows directly from the discrete-time argument used to derive (9') where $-C(t)dt$ is replaced by a general $dv(t)$ where $dv(t)$ is the instantaneous flow of funds from all non-capital gains sources.

It was necessary to derive (12) by starting with the discrete-time formulation because it is not obvious from the continuous version directly whether $dy - C(t)dt$ equals $\sum_1^n dN_i P_i + \sum_1^n dN_i dP_i$ or just $\sum_1^n dN_i P_i$.

⁽¹¹⁾ There are no other restrictions on the individual w_i because borrowing and short-selling are allowed.

where $m \equiv n-1$ and the w_1, \dots, w_m are unconstrained by virtue of the fact that the relation $w_n = 1 - \sum_1^m w_i$ will ensure that the identity constraint in (14) is satisfied.

4. Optimal portfolio and consumption rules: the equations of optimality. The problem of choosing optimal portfolio and consumption rules for an individual who lives T years is formulated as follows,

$$(15) \quad \text{Max } E_0 \left[\int_0^T U(C(t), t) dt + B(W(T), T) \right]$$

subject to: $W(0) = W_0$; the budget constraint (14), which in the case of a "risk-free" asset becomes (14'); and where the utility function (during life), U , is assumed to be strictly concave in C and the "bequest" function, B , is assumed also to be concave in W . ⁽¹²⁾

To derive the optimal rules, the technique of stochastic dynamic programming is used. Define

$$(16) \quad J(W, P, t) \equiv \text{Max}_{\{C, W\}} E_t \left[\int_t^T U(C, s) ds + B(W(T), T) \right]$$

where as before, " E_t " is the conditional expectation operator, conditional on $W(t) = W$ and $P_i(t) = P_i$. Define

$$(17) \quad \phi(w, C; W, P, t) \equiv U(C, t) + \mathcal{L}[J],$$

given $w_i(t) = w_i$, $C(t) = C$, $W(t) = W$, and $P_i(t) = P_i$. ⁽¹³⁾

⁽¹²⁾ Where there is no "risk-free" asset, it is assumed no asset can be expressed as a linear combination of the other assets, implying that the $n \times n$ variance-covariance matrix of returns, $\Omega = [\sigma_{ij}]$ where $\sigma_{ij} \equiv \rho_{ij} \sigma_i \sigma_j$, is non-singular. In the case when there is a "risk-free" asset, the same assumption is made about the "reduced" $m \times m$ variance-covariance matrix.

⁽¹³⁾ " \mathcal{L} " is short for the rigorous $\mathcal{L}_{P, W}^{W, C}$, the Dynkin operator over the variables P and W for a given set of controls w and C . $\mathcal{L} \equiv \frac{\partial}{\partial t} + \left[\sum_1^n w_i \alpha_i W - C \right] \frac{\partial}{\partial W} + \sum_1^n \alpha_i P_i \frac{\partial}{\partial P_i} + \frac{1}{2} \sum_1^n \sum_1^n \sigma_{ij} w_i w_j W^2 \frac{\partial^2}{\partial W^2} + \frac{1}{2} \sum_1^n \sum_1^n P_i P_j \sigma_{ij} \frac{\partial^2}{\partial P_i \partial P_j} + \sum_1^n \sum_1^n P_i W w_j \sigma_{ij} \frac{\partial^2}{\partial P_i \partial W}$

From the theory of stochastic dynamic programming, the following theorem provides the method for deriving the optimal rules, C^* and w^* .

Theorem I. ⁽¹⁴⁾ If the $P_i(t)$ are generated by a strong diffusion process, U is strictly concave in C , and B is concave in W , then there exists a set of optimal rules (controls), w^* and C^* , satisfying $\sum_1^n w_i^* = 1$ and $J(W, P, T) = B(W, T)$ and these controls satisfy

$$0 = \phi(C^*, w^*; W, P, t) \geq \phi(C, w; W, P, t) \\ \text{for } t \in [0, T].$$

From theorem I, we have that

$$(18) \quad 0 = \text{Max} \{ \phi(C, w; W, P, t) \} \\ \{C, w\}$$

In the usual fashion of maximization under constraint, we define the Lagrangian, $L \equiv \phi + \lambda[1 - \sum_1^n w_i]$ where λ is the multiplier and find the extreme points from the first-order conditions

$$(19) \quad 0 = L_C(C^*, w^*) = U_C(C^*, t) - J_W$$

$$(20) \quad 0 = L_{w_k}(C^*, w^*) = -\lambda + J_{W\alpha_k} W + J_{WW} \sum_1^n \sigma_{kj} w_j^* W^2 \\ + \sum_1^n J_{jW} \sigma_{kj} P_j W, \quad k=1, \dots, n$$

$$(21) \quad 0 = L_\lambda(C^*, w^*) = 1 - \sum_1^n w_i^*$$

where the notation for partial derivatives is $J_W \equiv \frac{\partial J}{\partial W}$, $J_t \equiv \frac{\partial J}{\partial t}$, $U_C \equiv \frac{\partial U}{\partial C}$, $J_i \equiv \frac{\partial J}{\partial P_i}$, $J_{ij} \equiv \frac{\partial^2 J}{\partial P_i \partial P_j}$, and $J_{jW} \equiv \frac{\partial^2 J}{\partial P_j \partial W}$.

⁽¹⁴⁾ For a heuristic proof of this theorem and the derivation of the stochastic Bellman equation, see Dreyfus [4] and Merton [12]. For a rigorous proof and discussion of weaker conditions, see Kushner [9], chapter IV especially theorem 7.

Because $L_{CC} = \phi_{CC} = U_{CC} < 0$, $L_{Cw_k} = \phi_{Cw_k} = 0$;

$L_{w_k w_k} = \sigma_k^2 W^2 J_{WW}$; $L_{w_k w_j} = 0$, $k \neq j$, a sufficient condition

for a unique interior maximum is that $J_{WW} < 0$ (i.e. that J be strictly concave in W). That assumed, as an immediate consequence of differentiating (19) totally with respect to W , we have

$$(22) \quad \frac{\partial C^*}{\partial W} > 0$$

To solve explicitly for C^* and w^* , we solve the $n+2$ non-dynamic implicit equations, (19) - (21), for C^* , and w^* , and λ as functions of J_W , J_{WW} , J_{jW} , W , P and t . Then C^* and w^* are substituted in (18) which now becomes a second-order partial differential equation for J , subject to the boundary condition $J(W, P, T) = B(W, T)$. Having (in principle at least) solved this equation for J , we then substitute back into (19) - (21) to derive the optimal rules as functions of W , P , and t . Define the inverse function $G \equiv [U_C]^{-1}$. Then from (19),

$$(23) \quad C^* = G(J_W, t)$$

To solve for the w_i^* , note that (20) is a linear system in w_i^* and hence can be solved explicitly. Define

$$(24) \quad \Omega \equiv [\sigma_{ij}] \text{ , the } nxn \text{ variance-covariance matrix}$$

$$[v_{ij}] \equiv \Omega^{-1} \quad (15)$$

$$\Gamma \equiv \sum_1^n \sum_1^n v_{ij} .$$

Eliminating λ from (20), the solution for w_k^* can be written as

$$(25) \quad w_k^* = h_k(P, t) + m(P, W, t)g_k(P, t) + f_k(P, W, t), \quad k=1, \dots, n$$

where $\sum_1^n h_k \equiv 1$, $\sum_1^n g_k \equiv 0$, and $\sum_1^n f_k \equiv 0$.⁽¹⁶⁾

⁽¹⁵⁾ Ω^{-1} exists by the assumption on Ω in footnote 12.

⁽¹⁶⁾ $h_k(P, t) \equiv \sum_1^n v_{kj} / \Gamma$; $m(P, W, t) \equiv -J_W / \Gamma W J_{WW}$;

(continued)

Substituting for w^* and C^* in (18), we arrive at the fundamental partial differential equation for J as a function of W , P , and t ,

$$\begin{aligned}
 (26) \quad 0 = & U[G, t] + J_t + J_W \left[\frac{\sum_1^n \sum_1^n v_{kj} \alpha_k^W}{\Gamma} - G \right] \\
 & + \sum_1^n J_i \alpha_i P_i + \frac{1}{2} \sum_1^n \sum_1^n J_{ij} \sigma_{ij} P_i P_j + \frac{W}{\Gamma} \sum_1^n J_{jW} P_j \\
 & - \frac{J_W}{\Gamma J_{WW}} \left(\sum_1^n \Gamma J_{kW} P_k \alpha_k - \sum_1^n J_{jW} P_j \sum_1^n \sum_1^n v_{kl} \alpha_l \right) \\
 & + \frac{J_{WW} W^2}{2\Gamma} - \frac{1}{2\Gamma J_{WW}} \left[\sum_1^n \sum_1^n J_{jW} J_{mW} P_j P_m \sigma_{mj} \Gamma - \left(\sum_1^n J_{iW} P_i \right)^2 \right] \\
 & - \frac{J_W^2}{2\Gamma J_{WW}} \left[\sum_1^n \sum_1^n v_{kl} \alpha_k \alpha_l \Gamma - \left(\sum_1^n \sum_1^n v_{kl} \alpha_k \right)^2 \right]
 \end{aligned}$$

subject to the boundary condition $J(W, P, T) = B(W, T)$. If (26) were solved, the solution J could be substituted into (23) and (25) to obtain C^* and w^* as functions of W , P , and t .

For the case where one of the assets is "risk-free", the equations are somewhat simplified because the problem can be solved directly as an unconstrained maximum by eliminating w_n as was done in (14'). In this case, the optimal proportions in the risky assets are

$$(27) \quad w_k^* = - \frac{J_W}{J_{WW} W} \sum_1^m v_{kj} (\alpha_j - r) - \frac{J_{kW} P_k}{J_{WW} W}, \quad k=1, \dots, m.$$

The partial differential equation for J corresponding to (26) becomes

$$\begin{aligned}
 (28) \quad 0 = & U[G, T] + J_t + J_W [rW - G] + \sum_1^m J_i \alpha_i P_i \\
 & + \frac{1}{2} \sum_1^m \sum_1^m J_{ij} \sigma_{ij} P_i P_j - \frac{J_W}{J_{WW}} \sum_1^m J_{jW} P_j (\alpha_j - r) \\
 & + \frac{J_W^2}{2J_{WW}} \sum_1^m \sum_1^m v_{ij} (\alpha_i - r) (\alpha_j - r) - \frac{1}{2J_{WW}} \sum_1^m \sum_1^m J_{iW} J_{jW} \sigma_{ij} P_i P_j
 \end{aligned}$$

(16) (continued)

$$\begin{aligned}
 g_k(P, t) \equiv & \sum_1^n v_{kl} \left(\Gamma \alpha_l - \sum_1^n \sum_1^n v_{ij} \alpha_j \right); \quad f_k(P, W, t) \equiv \left[\Gamma J_{kW} P_k \right. \\
 & \left. - \sum_1^n J_{iW} P_i \sum_1^n v_{kj} \right] / \Gamma W J_{WW}
 \end{aligned}$$

subject to the boundary condition $J(W,P,T) = B(W,T)$.

Although (28) is a simplified version of (26), neither (26) nor (28) lend themselves to easy solution. The complexities of (26) and (28) are caused by the basic non-linearity of the equations and the large number of state variables. Although there is little that can be done about the non-linearities, in some cases, it may be possible to reduce the number of state variables.

5. Log-normality of prices and the continuous-time analog to Tobin-Markowitz mean-variance analysis. When, for $k=1,\dots,n$, α_k and σ_k are constants, the asset prices have stationary, log-normal distributions. In this case, J will be a function of W and t only and not P . Then (26) reduces to

$$(29) \quad 0 = U[G,t] + J_t + J_W \left[\frac{\sum_1^n \sum_1^n v_{kj} \alpha_k}{\Gamma} W - G \right] + \frac{J_{WW} W^2}{2\Gamma} - \frac{J_W^2}{2\Gamma J_{WW}} \left[\sum_1^n \sum_1^n v_{kl} \alpha_k \alpha_l \Gamma - \left(\sum_1^n \sum_1^n v_{kl} \alpha_k \right)^2 \right]$$

From (25), the optimal portfolio rule becomes

$$(30) \quad w_k^* = h_k + m(W,t) g_k$$

where $\sum_1^n h_k \equiv 1$ and $\sum_1^n g_k \equiv 0$ and h_k and g_k are constants.

From (30), the following "separation" or "mutual fund" theorem can be proved.

Theorem II.⁽¹⁷⁾ Given n assets with prices P_i whose changes are stationary and log-normally distributed, then (1) there exist a unique pair (except for scale) of "mutual funds" constructed from linear combinations

⁽¹⁷⁾ See Cass and Stiglitz [1] for a general discussion of Separation theorems. The only degenerate case is when all the assets are identically distributed (i.e. symmetry) in which case, only one mutual fund is needed.

of these assets such that, independent of preferences (i.e. the form of the utility function) wealth distribution, or time horizon, individuals will be indifferent between choosing from a linear combination of these two funds or a linear combination of the original n assets. (2) If P_f is the price per share of either fund, then P_f is log-normally distributed. Further (3) if δ_k = percentage of one mutual fund's value held in the k^{th} asset and if λ_k = percentage of the other mutual fund's value held in the k^{th} asset, then one can find that

$$\delta_k = h_k + g_k, \quad k=1, \dots, n$$

$$\text{and } \lambda_k = h_k, \quad k=1, \dots, n.$$

Proof: (1) (30) is a parametric representation of a line in the hyperplane defined by $\sum_1^n w_k^* = 1$.⁽¹⁸⁾ Hence, there exist two linearly independent vectors (namely, the vectors of asset proportions held by the two mutual funds) which form a basis for all optimal portfolios chosen by the individuals. Therefore, each individual would be indifferent between choosing a linear combination of the mutual fund shares or a linear combination of the original n assets.

(2) Let $V \equiv N_f P_f$ = the total value of (either) fund where N_f = number of shares of the fund outstanding. Let N_k = number of shares of asset k held by the fund and $\mu_k \equiv N_k P_k / V$ = percentage of total value invested in the k^{th} asset. Then $V = \sum_1^n N_k P_k$ and

$$(31) \quad dV = \sum_1^n N_k dP_k + \sum_1^n P_k dN_k + \sum dP_k dN_k$$

$$= N_f dP_f + P_f dN_f + dP_f dN_f$$

⁽¹⁸⁾ See [1], page 15.

But

$$\begin{aligned}
 (32) \quad \sum_1^n P_k dN_k + \sum_1^n dP_k dN_k &= \text{net inflow of funds from non-} \\
 &\quad \text{capital gain sources} \\
 &= \text{net value of new shares issued} \\
 &= P_f dN_f + dN_f dP_f
 \end{aligned}$$

From (31) and (32), we have that

$$(33) \quad N_f dP_f = \sum_1^n N_k dP_k$$

By the definition of V and μ_k , (33) can be re-written as

$$\begin{aligned}
 (34) \quad \frac{dP_f}{P_f} &= \sum_1^n \mu_k \frac{dP_k}{P_k} \\
 &= \sum_1^n \mu_k \alpha_k dt + \sum_1^n \mu_k \sigma_k dz_k
 \end{aligned}$$

By Itô's Lemma and (34), we have that

$$\begin{aligned}
 (35) \quad P_f(t) &= P_f(0) \exp \left[\left(\sum_1^n \mu_k \alpha_k - \frac{1}{2} \sum_1^n \sum_1^n \mu_k \mu_j \sigma_{kj} \right) t \right. \\
 &\quad \left. + \sum_1^n \mu_k \sigma_k \int_0^t dz_k \right].
 \end{aligned}$$

So, $P_f(t)$ is log-normally distributed.

(3) Let $a(W,t;U) \equiv$ percentage of wealth invested in the first mutual fund by an individual with utility function U and wealth W at time t . Then, $(1-a)$ must equal the percentage of wealth invested in the second mutual fund. Because the individual is indifferent between these asset holdings or an optimal portfolio chosen from the original n assets, it must be that

$$(36) \quad w_k^* = h_k + m(W,t)g_k = a\delta_k + (1-a)\lambda_k, \quad k=1,\dots,n$$

Clearly, the only solution to the linear system (36) for all W , t , and U is

$$(37) \quad \delta_k = h_k + g_k, \quad k = 1, \dots, n$$

$$\lambda_k = h_k, \quad k = 1, \dots, n$$

$$a = m(W,t).$$

Note that $\sum_1^n \delta_k = \sum_1^n (h_k + g_k) \equiv 1$ and $\sum_1^n \lambda_k = \sum_1^n h_k \equiv 1$.
Q.E.D.

For the case when one of the assets is "risk-free", there is a corollary to theorem II. Namely,

Corollary. If one of the assets is "risk-free", then one of the two mutual funds will contain only this asset. If δ_k = percentage of the total value of the "risky" mutual fund invested in the k^{th} asset, then

$$\delta_k = \sum_1^m v_{kj}(\alpha_j - r) / \sum_1^m \sum_1^m v_{ij}(\alpha_j - r), \quad k=1, \dots, m.$$

Proof. By the assumption of stationary log-normal prices, (27) reduces to

$$(38) \quad w_k^* = \frac{-J_W}{J_{WW}} \sum_1^m v_{kj}(\alpha_j - r), \quad k=1, \dots, m$$

and

$$(39) \quad w_n^* = 1 - \sum_1^m w_k^* = 1 + \frac{J_W}{J_{WW}} \sum_1^m \sum_1^m v_{ij}(\alpha_j - r)$$

By the same argument as in the proof of theorem II, (38) and (39) define a line in the hyperplane defined by $\sum_1^n w_1^* = 1$ and

$$\delta_k = \sum_1^m v_{kj}(\alpha_j - r) / \sum_1^m \sum_1^m v_{ij}(\alpha_j - r); \quad \lambda_k = 0, \quad k=1, \dots, m$$

$$\delta_n = 0 \quad ; \quad \lambda_k = 1 \quad k=n.$$

Q.E.D.

Thus, if we have an economy where all asset prices are log-normally distributed, the investment decision can be divided into two parts by the establishment of two financial intermediaries (mutual funds) to hold all individual securities and to issue shares of their own for purchase by individual investors. The separation is complete because the "instructions" given the fund managers, namely to hold pro-

portions δ_k and λ_k of the k^{th} security, $k=1, \dots, n$, depend only on the price distribution parameters and are independent of individual preferences, wealth distribution, or age distribution.

The similarity of this result to that of the classical Tobin-Markowitz analysis is clearest when one examines closely the investment rule given to the "risky" fund's manager when there exists a "risk-free" asset (money) with zero return ($r=0$). It is easy to show that the δ_k proportions prescribed in the corollary are derived by finding the locus of points in the (instantaneous) mean-standard deviation space of composite returns which minimize variance for a given mean, and then by finding the point where a line drawn from the origin is tangent to the locus. This point determines the δ_k as is illustrated in figure 1.

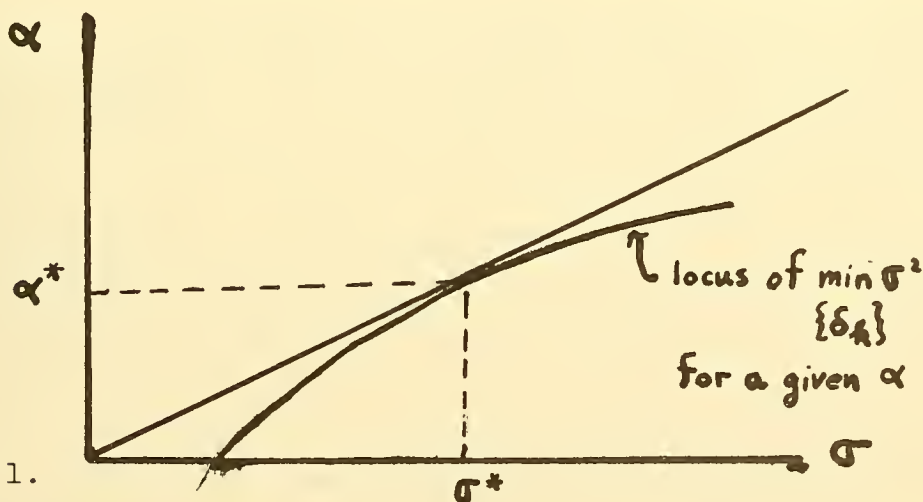


figure 1.

Given the α^* , the δ_k are determined. So the log-normal assumption in the continuous-time model is sufficient to allow the same analysis as in the static mean-variance model but without the objectionable assumptions of quadratic utility or normality of the distribution of absolute price changes. (Log-normality of price changes is much less objectionable, since this does invoke "limited liability" and by the central limit theorem is the only regular solution to

any continuous-space, infinitely-divisible process in time.)

An immediate advantage for the present analysis is that whenever log-normality of prices is assumed, we can work, without loss of generality, with just two assets, one "risk-free" and one risky with its price log-normally distributed. The risky asset can always be thought of as a composite asset with price $P(t)$ defined by the process

$$(40) \quad \frac{dP}{P} = \alpha dt + \sigma dz$$

where

$$(41) \quad \alpha \equiv \sum_1^m \sum_1^m v_{kj} (\alpha_j - r) \alpha_k / \sum_1^m \sum_1^m v_{ij} (\alpha_j - r)$$

$$\sigma^2 \equiv \sum_1^m \sum_1^m \delta_k \delta_j \sigma_{kj}$$

$$dz \equiv \sum_1^m \delta_k \sigma_k dz_k / \sigma$$

6. Explicit solutions for a particular class of utility functions. On the assumption of log-normality of prices, some characteristics of the asset demand functions were shown. If a further assumption about the preferences of the individual is made, then equation (28) can be solved in closed form, and the optimal consumption and portfolio rules derived explicitly. Assume that the utility function for the individual, $U(C,t)$, can be written as $U(C,t) = e^{-\rho t} V(C)$ where V is a member of the family of utility functions whose measure of absolute risk aversion is positive and hyperbolic in consumption, i.e. $A(C) \equiv -V''/V' = 1 / \left(\frac{C}{1-\gamma} + \eta/\beta \right) > 0$, subject to the restrictions

$$(42) \quad \gamma \neq 1; \beta > 0; \left(\frac{\beta C}{1-\gamma} + \eta \right) > 0; \eta = 1 \text{ if } \gamma = 1.$$

All members of the HARA (hyperbolic absolute risk-aversion) family can be expressed as

$$(43) \quad V(C) = \frac{(1-\gamma)}{\gamma} \left(\frac{\beta C}{1-\gamma} + \eta \right)^{\gamma}$$

This family is rich, in the sense that by suitable adjustment of the parameters, one can have a utility function with absolute or relative risk aversion increasing, decreasing, or constant. (19)

(19)

table 1. Properties of HARA Utility Functions

$$A(C) = \frac{1}{\frac{C}{1-\gamma} + \frac{\eta}{\beta}} > 0 \quad (\text{implies } \eta > 0 \text{ for } \gamma > 1)$$

$$A'(C) = \frac{-1}{(1-\gamma) \left(\frac{C}{1-\gamma} + \frac{\eta}{\beta} \right)^2} \begin{aligned} &< 0 \text{ for } -\infty < \gamma < 1 \\ &> 0 \text{ for } 1 < \gamma < \infty \\ &= 0 \text{ for } \gamma = \pm \infty \end{aligned}$$

$$\text{Relative risk aversion } R(C) \equiv -V''C/V' = A(C)C$$

$$R'(C) = \frac{\eta/\beta}{\left(\frac{C}{1-\gamma} + \frac{\eta}{\beta} \right)^2} \begin{aligned} &> 0 \text{ for } \eta > 0 \quad (-\infty \leq \gamma \leq \infty, \gamma \neq 1) \\ &= 0 \text{ for } \eta = 0 \\ &< 0 \text{ for } \eta < 0 \quad (-\infty < \gamma < 1) \end{aligned}$$

Note that included as members of the HARA family are the widely-used iso-elastic (constant relative risk aversion), exponential (constant absolute risk aversion), and quadratic utility functions. As is well known for the quadratic case, the members of the HARA family with $\gamma > 1$ are only defined for a restricted range of consumption, namely $0 < C < (\gamma - 1)\eta/\beta$. [6], [10], [5], [13], [1], and [12] discuss the properties of various members of the HARA family in a portfolio context. Although this is not done here, the HARA definition can be generalized to include the cases when γ , β , and η are functions of time subject to the restrictions in (42).

Without loss of generality, assume that there are two assets, one "risk-free" asset with return r and the other, a "risky" asset whose price is log-normally distributed satisfying (40). From (28), the optimality equation for J is

$$(44) \quad 0 = \frac{(1-\gamma)^2}{\gamma} e^{-\rho t} \left[\frac{e^{\rho t} J_W}{\beta} \right]^{\frac{\gamma}{\gamma-1}} + J_t + \left[(1-\gamma)\eta/\beta + rW \right] J_W - \frac{J_W^2}{J_{WW}} \frac{(\alpha-r)^2}{\alpha\sigma^2}$$

subject to $J(W,T) = 0$.⁽²⁰⁾ The equations for the optimal consumption and portfolio rules are

$$(45) \quad C^*(t) = \frac{(1-\gamma)}{\beta} \left[\frac{e^{\rho t} J_W}{\beta} \right]^{\frac{1}{\gamma-1}} - \frac{(1-\gamma)\eta}{\beta}$$

and

$$(46) \quad w^*(t) = \frac{-J_W}{J_{WW}W} \frac{(\alpha-r)}{\sigma^2}$$

where $w^*(t)$ is the optimal proportion of wealth invested in the risky asset at time t . A solution⁽²¹⁾ to (44) is

$$(47) \quad J(W,t) = \delta \beta^{-\gamma} e^{-\rho t} \left[\frac{\delta \left(1 - e^{-\left(\frac{\rho - \gamma v}{\delta} \right) (T-t)} \right)}{\delta - \gamma v} \right]^{\delta} \left[\frac{W}{\delta} + \frac{\eta}{\beta r} \left(1 - e^{-r(T-t)} \right) \right]^{\gamma}$$

where $\delta \equiv 1 - \gamma$ and $v \equiv r + (\alpha-r)^2/2\delta\sigma^2$.

⁽²⁰⁾ It is assumed for simplicity that the individual has a zero bequest function, i.e. $B \equiv 0$. If $B(W,T) \equiv H(T) (aW+b)^{\gamma}$, the basic functional form for J in (47) will be the same. Otherwise, systematic effects of age will be involved in the solution.

⁽²¹⁾ By theorem I, there is no need to be concerned with uniqueness although, in this case, the solution is unique.

From (45), (46), and (47), the optimal consumption and portfolio rules can be written in explicit form as

$$(48) \quad C^*(t) = \frac{[\rho - \gamma v] \left[W(t) + \frac{\delta \eta}{\beta r} (1 - e^{r(t-T)}) \right]}{\delta \left(1 - \exp \left[\frac{(\rho - \gamma v)}{\delta} (t-T) \right] \right)} - \frac{\delta \eta}{\beta}$$

and

$$(49) \quad w^*(t)W(t) = \frac{(\alpha - r)}{\delta \sigma^2} W(t) + \frac{\eta(\alpha - r)}{\beta r \sigma^2} (1 - e^{r(t-T)})$$

The manifest characteristic of (48) and (49) is that the demand functions are linear in wealth. It will be shown that the HARA family is the only class of concave utility functions which imply linear solutions. For notation purposes, define $I(X, t) \in \text{HARA}(X)$ if $-I_{XX}/I_X = 1/(\alpha X + \beta) > 0$, where α and β are, at most, functions of time and I is a strictly concave function of X .

Theorem III. Given the model specified in this section, then $C^* = aW + b$ and $w^*W = gW + h$ where a , b , g , and h are, at most, functions of time if and only if $U(C, t) \in \text{HARA}(C)$.

Proof: "If" part is proved directly by (48) and (49).

"Only if" part: Suppose $w^*W = gW + h$ and $C^* = aW + b$

From (19), we have that $U_C(C^*, t) = J_W(W, t)$. Differentiating this expression totally with respect to W , we have that $U_{CC} \frac{dC^*}{dW} = J_{WW}$ or $aU_{CC} = J_{WW}$ and hence

$$(50) \quad \frac{-U_{CC}a}{U_C} = -\frac{J_{WW}}{J_W}.$$

From (46), $w^*W = gW + h = -J_W(\alpha - r)/J_{WW}\sigma^2$ or

$$(51) \quad -J_{WW}/J_W = 1 / \left[\left(\frac{\sigma^2 g}{(\alpha - r)} \right) W + \frac{\sigma^2 h}{(\alpha - r)} \right].$$

So, from (50) and (51), we have that U must satisfy

$$(52) \quad -U_{CC}/U_C = 1/(a'c^* + b')$$

where $a' \equiv \sigma^2 g/(\alpha-r)$ and $b' \equiv (\alpha\sigma^2 h - b\sigma^2 g)/(\alpha-r)$.

Hence $U \in \text{HARA}(C)$. Q.E.D.

As an immediate result of theorem III, a second theorem can be proved.

Theorem IV. Given the model specified in this section, $J(W,t) \in \text{HARA}(W)$ if and only if $U \in \text{HARA}(C)$.

Proof: "If" part is proved directly by (47).

"Only if" part: suppose $J(W,t) \in \text{HARA}(W)$. Then from (46), w^*W is a linear function of W . If (28) is differentiated totally with respect to wealth and given the specific price behavior assumptions of this section, we have that C^* must satisfy

$$(53) \quad C^* = rW + \frac{J_{tW}}{J_{WW}} + \frac{rJ_W}{J_{WW}} - w^*W \frac{d(w^*W)}{dW} - \frac{J_{WWW}}{2J_{WW}} \left(\frac{J_W}{J_{WW}} \right)^2 \frac{(\alpha-r)^2}{\sigma^2}$$

But if $J \in \text{HARA}(W)$, then (53) implies that C^* is linear in wealth. Hence, by theorem III, $U \in \text{HARA}(C)$. Q.E.D.

Given (48) and (49), the stochastic process which generates wealth when the optimal rules are applied, can be derived.

From the budget equation (14'), we have that

$$(54) \quad \begin{aligned} dW &= \left\{ \left[(w^*(\alpha-r) + r)W - C^* \right] dt + \sigma w^*W dz \right. \\ &= \left\{ \left[\frac{(\alpha-r)^2}{\sigma^2 \delta} - \frac{\mu}{1-e^{\mu(t-T)}} \right] dt + \frac{(\alpha-r)}{\sigma \delta} dz \right\} X(t) \\ &\quad + r \left[W + \frac{\delta \eta}{\beta r} \right] dt \end{aligned}$$

where $X(t) \equiv W(t) + \frac{\delta \eta}{\beta r} (1-e^{r(t-T)})$ for $0 \leq t \leq T$ and

$\mu \equiv (\rho-v)/\delta$. By Itô's Lemma, $X(t)$ is the solution to

$$(55) \quad \frac{dX}{X} = \left[\delta - \frac{\mu}{(1-e^{\mu(t-T)})} \right] dt + \frac{(\alpha-r)}{\sigma \delta} dz$$

Again using Itô's Lemma, integrating (55) we have that

$$(56) \quad X(t) = X(0) \exp \left\{ \left[\delta - \mu - \frac{(\alpha - r)^2}{2\sigma^2 \delta^2} \right] t + \frac{(\alpha - r)}{\sigma \delta} \int_0^t dz \right\} (1 - e^{\mu(t-T)}) / (1 - e^{-\mu T})$$

and hence, $X(t)$ is log-normally distributed. Therefore,

$W(t) = X(t) - \frac{\delta \eta}{\beta r} (1 - e^{r(t-T)})$ is a "displaced" or "three-parameter" log-normally distributed random variable. By Itô's Lemma, solution (56) to (55) holds with probability one and because $W(t)$ is a continuous process, we have with probability one that

$$(57) \quad \lim_{t \rightarrow T} W(t) = 0.$$

From (48), with probability one,

$$(58) \quad \lim_{t \rightarrow T} C^*(t) = 0$$

Further, from (48), $C^* + \frac{\delta \eta}{\beta}$ is proportional to $X(t)$ and from the definition of $U(C^*, t)$, $U(C^*, t)$ is a log-normally distributed random variable.⁽²²⁾ The following theorem shows that this result holds only if $U(C, t) \in \text{HARA}(C)$.

Theorem V. Given the model specified in this section and the time-dependent random variable $Y(t) \equiv U(C^*, t)$, then Y is log-normally distributed if and only if $U(C, t) \in \text{HARA}(C)$.

Proof: "If" part: It was previously shown that if $U \in \text{HARA}(C)$ then Y is log-normally distributed.

"only if" part: Let $C^* \equiv g(W, t)$ and $w^*W \equiv f(W, t)$. By Itô's Lemma,

$$(59) \quad \begin{aligned} dY &= U_C dC^* + U_t dt + \frac{1}{2} U_{CC} (dC^*)^2 \\ dC^* &= g_W dW + g_t dt + \frac{1}{2} g_{WW} (dW)^2 \\ dW &= [f(\alpha - r) + rW - g] dt + \sigma dz \end{aligned}$$

⁽²²⁾ $U = \frac{(1-\gamma)}{\gamma} e^{-\rho t} \left[\frac{gC}{1-\gamma} + \eta \right]^\gamma$ and products and powers of log-normal variates are log normal with one exception: the logarithmic utility function ($\gamma=0$) is a singular case where $U(C^*, t) = \log C^*$ is normally distributed.

Because $(dW)^2 = \sigma^2 f^2 dt$, we have that

$$(60) \quad dC^* = \left[g_W f(\alpha - r) + g_W r W - g g_W + \frac{1}{2} g_{WW} \sigma^2 f^2 + g_t \right] dt + \sigma f g_W dz$$

and

$$(61) \quad dY = \left\{ U_C \left[g_W f(\alpha - r) + r g_W W - g g_W + \frac{1}{2} g_{WW} \sigma^2 f^2 + g_t \right] + U_t + \frac{1}{2} U_{CC} \sigma^2 f^2 g_W^2 \right\} dt + \sigma f g_W U_C dz$$

A necessary condition for Y to be log-normal is that Y satisfy.

$$(62) \quad \frac{dY}{Y} = F(Y)dt + b dz$$

where b is, at most, a function of time. If Y is log-normal, from (61) and (62), we have that

$$(63) \quad b(t) = \sigma f g_W U_C / U$$

From the first-order conditions, f and g must satisfy

$$(64) \quad U_{CC} g_W = J_{WW}$$

$$f = - J_W(\alpha - r) / \sigma^2 J_{WW}$$

But (63) and (64) imply that

$$(65) \quad b U / \sigma U_C = f g_W = -(\alpha - r) U_C / \sigma^2 U_{CC}$$

or

$$(66) \quad - U_{CC} / U_C = \eta(t) U_C / U$$

where $\eta(t) \equiv (\alpha - r) / \sigma b(t)$. Integrating (66), we have that

$$(67) \quad U = [(\eta + 1) (C + \mu) \zeta(t)]^{\frac{1}{\eta + 1}}$$

where $\zeta(t)$ and μ are, at most, functions of time and hence $U \in \text{HARA}(C)$. Q.E.D.

For the case when asset prices satisfy the "geometric" Brownian motion hypothesis and the individual's utility function is a member of the HARA family, the consumption-portfolio problem is completely solved. From (48) and (49), one could examine the effects of shifts in various parameters on the consumption and portfolio rules by the methods of comparative statics as was done for the iso-elastic case in [12].

7. Non-capital gains income:wages. In the previous sections, it was assumed that all income was generated by capital gains. If a (certain) wage income flow, $dy = Y(t)dt$, is introduced, the optimality equation (18) becomes

$$(68) \quad 0 = \text{Max}_{\{C, w\}} [U(C, t) + \bar{\mathcal{L}}(J)]$$

where the operator $\bar{\mathcal{L}}$ is defined by $\bar{\mathcal{L}} \equiv \mathcal{L} + Y(t) \frac{\partial}{\partial W}$.

This new complication causes no particular computational difficulties. If a new control variable, $\tilde{C}(t)$, and new utility function, $V(C, t)$ are defined by $\tilde{C}(t) \equiv C(t) - Y(t)$ and $V(\tilde{C}, t) \equiv U(\tilde{C}(t) + Y(t), t)$, then (68) can be re-written as

$$(69) \quad 0 = \text{Max}_{\{C, w\}} [V(C, t) + \mathcal{L}[J]]$$

which is the same equation as the optimality equation (18) when there is no wage income and where consumption has been re-defined as consumption in excess of wage income.

In particular, if $Y(t) \equiv Y$, a constant, and $U \in \text{HARA}(C)$, then the optimal consumption and portfolio rules corresponding to (48) and (49) are

$$(70) \quad C^*(t) = \frac{[\rho - \gamma v]}{\delta} \left[W + \frac{Y(1 - e^{r(t-T)})}{r} + \frac{\delta \eta}{\beta r} (1 - e^{r(t-T)}) \right] - \frac{\delta \eta}{\beta}$$

and

$$(71) \quad w^*W = \frac{(\alpha - r)}{\delta \sigma^2} \left(W + \frac{Y(1 - e^{r(t-T)})}{r} \right) + \frac{(\alpha - r)\eta}{\beta r \sigma^2} (1 - e^{r(t-T)})$$

Comparing (70) and (71) with (48) and (49), one finds that, in computing the optimal decision rules, the individual capitalizes the lifetime flow of wage income at the market (risk-free) rate of interest and then treats the capitalized value

as an addition to the current stock of wealth.⁽²³⁾

The introduction of a stochastic wage income will cause increased computational difficulties although the basic analysis is the same as for the no-wage income case. For a solution to a particular example of a stochastic wage problem, see example two of section eight.

8. Poisson processes. The previous analyses always assumed that the underlying stochastic processes were smooth functions of Brownian motions and therefore, continuous in both the time and state spaces. Although such processes are reasonable models for price behavior of many types of liquid assets, they are rather poor models for the description of other types. The Poisson process is a continuous-time process which allows discrete (or dis-continuous) changes in the variables. The simplest independent Poisson process defines the probability of an event occurring during a time interval of length h (where h is as small as you like) as follows,

$$(72) \quad \begin{aligned} \text{Prob \{the event does not occur in the time interval } (t, t+h)\} \\ = 1 - \lambda h + O(h) \end{aligned}$$

$$\begin{aligned} \text{Prob \{the event occurs once in the time interval } (t, t+h)\} \\ = \lambda h + O(h) \end{aligned}$$

$$\begin{aligned} \text{Prob \{the event occurs more than once in the time interval} \\ (t, t+h)\} = O(h) \end{aligned}$$

where $O(h)$ is the asymptotic order symbol defined by

$$(73) \quad \psi(h) \text{ is } O(h) \text{ if } \lim_{h \rightarrow 0} (\psi(h)/h) = 0$$

and λ = the mean number of occurrences per unit time.

Given the Poisson process, the "event" can be defined in a number of interesting ways. To illustrate the degree of

⁽²³⁾ As Hakansson [6] has pointed out, (70) and (71) are consistent with the Friedman Permanent Income and the Modigliani Life-cycle hypotheses. However, in general, this result will not hold.

latitude, three examples of applications of Poisson processes in the consumption-portfolio choice problem are presented below. Before examining these examples, it is first necessary to develop some of the mathematical properties of Poisson processes. There is a theory of stochastic differential equations for Poisson processes similar to the one for Brownian motion discussed in section two. Let $q(t)$ be an independent Poisson process with probability structure as described in (72). Let the event be that a state variable $x(t)$ has a jump in amplitude of size \mathcal{J} where \mathcal{J} is a random variable whose probability measure has compact support. Then a Poisson differential equation for $x(t)$ can be written as

$$(74) \quad dx = f(x,t)dt + g(x,t)dq$$

and the corresponding differential generator, \mathcal{L}_x , is defined by

$$(75) \quad \mathcal{L}_x[h(x,t)] \equiv h_t + f(x,t)h_x + E_t \{ \lambda [h(x + \mathcal{J}g,t) - h(x,t)] \}$$

where " E_t " is the conditional expectation over the random variable \mathcal{J} , conditional on knowing $x(t) = x$, and where $h(x,t)$ is a C^1 function of x and t .⁽²⁴⁾ Further, theorem I holds for Poisson processes.⁽²⁵⁾

Returning to the consumption-portfolio problem, consider first the two-asset case. Assume that one asset is a common stock whose price is log-normally distributed and that the other asset is a "risky" bond which pays an instantaneous rate of interest r when not in default but in the event of default, the price of the bond becomes zero.⁽²⁶⁾

⁽²⁴⁾ For a short discussion of Poisson differential equations and a proof of (75) as well as other references, see Kushner [9], pages 18-22.

⁽²⁵⁾ See Dreyfus [4], p.225 and Kushner [9], chapter IV.

⁽²⁶⁾ That the price of the bond is zero in the event of default is an extreme assumption made only to illustrate how default can be treated in the analysis. One could make the more reasonable assumption that the price in the event of default is a random variable. The degree of computational difficulty caused by this more reasonable assumption will depend on the choice of distribution for the random variable as well as the utility function of the individual.

From (74), the process which generates the bond's price can be written as

$$(76) \quad dP = rPdt - Pdq$$

where dq is as previously defined and $\mathcal{J} \equiv 1$ with probability one. Substituting the explicit price dynamics into (14'), the budget equation becomes

$$(77) \quad dW = \{wW(\alpha-r) + rW-C\}dt + w\sigma Wdz - (1-w)Wdq$$

From (75), (77), and theorem I, we have that the optimality equation can be written as

$$(78) \quad 0 = U(C^*,t) + J_t(W,t) + \lambda[J(w^*W,t) - J(W,t)] \\ + J_W(W,t) \left[(w^*(\alpha-r) + r)W - C^* \right] + \frac{1}{2} J_{WW}(W,t) \sigma^2 w^{*2} W^2$$

where C^* and w^* are determined by the implicit equations

$$(79) \quad 0 = U_C(C^*,t) - J_W(W,t)$$

and

$$(80) \quad 0 = \lambda J_W(w^*W,t) + J_W(W,t)(\alpha-r) + J_{WW}(W,t) \sigma^2 w^*W.$$

To see the effect of default on the portfolio and consumption decisions, consider the particular case when $U(C,t) \equiv C^\gamma/\gamma$, for $\gamma < 1$. The solutions to (79) and (80) are

$$(79') \quad C^*(t) = AW(t)/(1-\gamma) \left(1 - \exp[A(t-T)/(1-\gamma)] \right)$$

$$\text{where } A \equiv -\gamma \left[\frac{(\alpha-r)^2}{2\sigma^2(1-\gamma)} + r \right] + \lambda \left[1 - \frac{(2-\gamma)w^*\gamma}{\gamma} - \frac{\gamma(\alpha-r)}{2\sigma^2(1-\gamma)} w^* \gamma^{-1} \right]$$

and

$$(80') \quad w^* = \frac{(\alpha-r)}{\sigma^2(1-\gamma)} + \frac{\lambda}{\sigma^2(1-\gamma)} (w^*)^{\gamma-1} \quad (27')$$

As might be expected, the demand for the common stock is an increasing function of λ and, for $\lambda > 0$, $w^* > 0$ holds for all values of α, r , or σ^2 .

(27') Note that (79') and (80') with $\lambda = 0$ reduce to the solutions (48) and (49) when $\eta = \rho = 0$ and $\beta = 1 - \gamma$.

For the second example, consider an individual who receives a wage, $Y(t)$, which is incremented by a constant amount ϵ at random points in time. Suppose that the event of a wage increase is distributed Poisson with parameter λ . Then, the dynamics of the wage-rate state variable are described by

$$(81) \quad dY = \epsilon dq, \text{ with } \mathcal{J} \equiv 1 \text{ with probability one.}$$

Suppose further that the individual's utility function is of the form $U(C,t) \equiv e^{-\rho t} V(C)$ and that his time horizon is infinite (i.e. $T = \infty$).⁽²⁸⁾ Then, for the two-asset case of section six, the optimality equation can be written as

$$(83) \quad 0 = V(C^*) - \rho I(W,Y) + \lambda [I(W,Y + \epsilon) - I(W,Y)] \\ + I_W(W,Y) \left[(w^*(\alpha - r) + r)W + Y - C^* \right] + \frac{1}{2} I_{WW}(W,Y) \sigma^2 w^{*2} W^2$$

where $I(W,Y) \equiv e^{\rho t} J(W,Y,t)$. If it is further assumed that $V(C) = -e^{-\eta C}/\eta$, then the optimal consumption and portfolio rules, derived from (83), are

$$(84) \quad C^*(t) = r \left[W(t) + \frac{Y(t)}{r} + \frac{\lambda}{r^2} \left(\frac{1 - e^{-\eta \epsilon}}{\eta} \right) \right] + \frac{1}{\eta r} \left[\rho - r + \frac{(\alpha - r)^2}{2\sigma^2} \right]$$

and

$$(85) \quad w^*(t)W(t) = \frac{(\alpha - r)}{\eta \sigma^2 r}$$

In (84), $\left[W(t) + Y(t)/r + \lambda (1 - e^{-\eta \epsilon})/\eta r^2 \right]$ is the general wealth term, equal to the sum of present wealth and capitalized future wage earnings. If $\lambda = 0$, then (84) reduces to (70) in

⁽²⁸⁾ I have shown elsewhere [[12], p.252] that if $U = e^{-\rho t} V(C)$ and U is bounded or ρ sufficiently large to ensure convergence of the integral and if the underlying stochastic processes are stationary, then the optimality equation (18) can be written, independent of explicit time, as

$$(82) \quad 0 = \text{Max}_{\{C,w\}} [V(C) + \tilde{\mathcal{L}} [I]]$$

where $\tilde{\mathcal{L}} \equiv \mathcal{L} - \rho - \frac{\partial}{\partial t}$ and $I(W,P) \equiv e^{\rho t} J(W,P,t)$.

A solution to (82) is called the "stationary" solution to the consumption-portfolio problem. Because the time state variable is eliminated, solutions to (82) are computationally easier to find than for the finite-horizon case.

section seven, where the wage rate was fixed and known with certainty. When $\lambda > 0$, $\lambda(1-e^{-\eta\epsilon})/\eta r^2$ is the capitalized value of (expected) future increments to the wage rate, capitalized at a somewhat higher rate than the risk-free market rate reflecting the risk-aversion of the individual.⁽²⁹⁾

Let $X(t)$ be the "Certainty-equivalent wage rate at time t " defined as the solution to

$$(86) \quad U[X(t)] = E_0 U[Y(t)].$$

For this example, $X(t)$ is calculated as follows

$$\begin{aligned} (87) \quad -\frac{e^{-\eta X(t)}}{\eta} &= -\frac{1}{\eta} E_0 e^{-\eta Y(t)} \\ &= -\frac{1}{\eta} e^{-\eta Y(0)} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} e^{-\eta k \epsilon} \\ &= -\frac{1}{\eta} e^{-\eta Y(0) - \lambda t + \lambda e^{-\eta \epsilon} t} \end{aligned}$$

Solving for $X(t)$ from (87), we have that

$$(88) \quad X(t) = Y(0) + \lambda t (1 - e^{-\eta \epsilon}) / \eta.$$

The capitalized value of the (certainty equivalent wage income flow is

$$\begin{aligned} (89) \quad \int_0^{\infty} e^{-rs} X(s) ds &= \int_0^{\infty} Y(0) e^{-rs} ds + \int_0^{\infty} \frac{\lambda(1-e^{-\eta \epsilon})}{\eta} s e^{-rs} ds \\ &= \frac{Y(0)}{r} + \frac{\lambda(1-e^{-\eta \epsilon})}{\eta r^2} \end{aligned}$$

Thus, for this example,⁽³⁰⁾ the individual, in computing the

⁽²⁹⁾ The usual expected present discounted value of the increments to the wage flow is $E_t \int_t^{\infty} e^{-r(s-t)} [Y(s) - Y(t)] ds$
 $= \int_t^{\infty} \lambda \epsilon e^{-r(s-t)} (s-t) ds = \lambda \epsilon / r^2$, which is greater than $\lambda(1-e^{-\eta \epsilon})/\eta r^2$ for $\epsilon > 0$.

⁽³⁰⁾ The reader should not infer that this result holds in general. Although (86) is a common definition of Certainty-equivalent in one-period utility-of-wealth models, it is not satisfactory for dynamic consumption-portfolio models. The reason it works for this example is due to the particular relationship between the J and U functions when U is exponential.

present value of future earnings, determines the Certainty-equivalent flow and then capitalizes this flow at the (certain) market rate of interest.

The third example of a Poisson process differs from the first two because the occurrence of the event does not involve an explicit change in a state variable. Consider an individual whose age of death is a random variable. Further assume that the event of death at each instant of time is an independent Poisson process with parameter λ . Then, the age of death, τ , is the first time that the event (of death) occurs and is an exponentially distributed random variable with parameter λ . The optimality criterion is to

$$(90) \quad \text{Max } E_0 \{ \int_0^\tau U(C,t)dt + B(W(\tau),\tau) \}$$

and the associated optimality equation is

$$(91) \quad 0 = U(C^*,t) + \lambda[B(W,t) - J(W,t)] + \mathcal{L}[J].$$

To derive (91), an "artificial" state variable, $x(t)$, is constructed with $x(t) = 0$ while the individual is alive and $x(t) = 1$ in the event of death. Therefore, the stochastic process which generates x is defined by

$$(92) \quad dx = dq \text{ and } \mathcal{L} \equiv 1 \text{ with probability one}$$

and τ is now defined by x as

$$(93) \quad \tau = \min \{t | t > 0 \text{ and } x(t) = 1\}.$$

The derived utility function, J , can be considered a function of the state variables W , x , and t subject to the boundary condition

$$(94) \quad J(W,x,t) = B(W,t) \text{ when } x = 1.$$

In this form, example three is shown to be of the same type as examples one and two in that the occurrence of the Poisson event causes a state variable to be incremented, and (91) is of the same form as (78) and (83).

A comparison of (91) for the particular case when $B \equiv 0$ (no bequests) with (82) suggested the following theorem.⁽³¹⁾

⁽³¹⁾ I believe that a similar theorem has been proved by
(continued)

Theorem VI. If τ is as defined in (93) and U is such that the integral $E_0[\int_0^\tau U(C,t)dt]$ is absolutely convergent, then the maximization of $E_0[\int_0^\tau U(C,t)dt]$ is equivalent to the maximization of $\mathcal{E}_0[\int_0^\infty e^{-\lambda t} U(C,t)dt]$ where " E_0 " is the conditional expectation operator over all random variables including τ and " \mathcal{E}_0 " is the conditional expectation operator over all random variables excluding τ .

Proof: τ is distributed exponentially and is independent of the other random variables in the problem. Hence, we have that

$$(95) \quad E_0[\int_0^\tau U(C,t)dt] = \int_0^\infty \lambda e^{-\lambda \tau} \mathcal{E}_0[\int_0^\tau U(C,t)dt] d\tau \\ = \int_0^\infty \int_0^\tau \lambda g(t) e^{-\lambda \tau} dt d\tau$$

where $g(t) \equiv \mathcal{E}_0[U(C,t)]$. Because the integral in (95) is absolutely convergent, the order of integration can be interchanged, i.e. $\mathcal{E}_0[\int_0^\tau U(C,t)dt] = \int_0^\tau \mathcal{E}_0[U(C,t)]dt$. By integration by parts, (95) can be re-written as

$$(96) \quad \int_0^\infty \int_0^\tau e^{-\lambda \tau} g(t) dt d\tau = \int_0^\infty e^{-\lambda s} g(s) ds \\ = \mathcal{E}_0[\int_0^\infty e^{-\lambda t} U(C,t)dt]. \quad \text{Q.E.D.}$$

Thus, an individual who faces an exponentially-distributed uncertain age of death acts as if he will live forever, but with a subjective rate of time preference equal to his "force of mortality", i.e. to the reciprocal of his life expectancy.

9. Alternative price expectations to the geometric Brownian motion. The assumption of the geometric Brownian motion hypothesis is a rich one because it is a reasonably good

(31) (continued)

J.A. Mirrlees, but I have no reference. D. Cass and M.E. Yaari, in "Individual Saving, Aggregate Capital Accumulation, and Efficient Growth", in Essays on the Theory of Optimal Economic Growth, ed. K. Shell, (M.I.T. Press 1967), prove a similar theorem on page 262.

model of observed stock price behavior and it allows the proof of a number of strong theorems about the optimal consumption-portfolio rules, as was illustrated in the previous sections. However, as mentioned in the Introduction, there have been some disagreements with the underlying assumptions required to accept this hypothesis. The geometric Brownian motion hypothesis best describes a stationary equilibrium economy where expectations about future returns have settled down, and as such, really describes a "long-run" equilibrium model for asset prices. Therefore, to explain "short-run" consumption and portfolio selection behavior one must introduce alternative models of price behavior which reflect the dynamic adjustment of expectations.

In this section, alternative price behavior mechanisms are postulated which attempt to capture in a simple fashion the effects of changing expectations, and then comparisons are made between the optimal decision rules derived under these mechanisms with the ones derived in the previous sections. The choices of mechanisms are not exhaustive nor are they necessarily representative of observed asset price behavior. Rather they have been chosen as representative examples of price adjustment mechanisms commonly used in economic and financial models.

Little can be said in general about the form of a solution to (28) when α_k and σ_k depend in an arbitrary manner on the price levels. If it is specified that the utility function is a member of the HARA family, i.e.

$$(97) \quad U(C,t) = \frac{(1-\gamma)}{\gamma} F(t) \left(\frac{\beta C}{1-\gamma} + \eta \right)^\gamma$$

subject to the restrictions in (42), then (28) can be simplified because $J(W,P,t)$ is separable into a product of functions, one depending on W and t , and the other on P and t .⁽³²⁾

⁽³²⁾ This separability property was noted in [1], [5], [6], [10], [12], and [13]. It is assumed throughout this section that the bequest function satisfies the conditions of footnote 20.

In particular, if we take $J(W, P, t)$ to be of the form

$$(98) \quad J(W, P, t) = \frac{(1-\gamma)}{\gamma} H(P, t) F(t) \left(\frac{W}{1-\gamma} + \frac{\eta}{\beta r} [1 - e^{r(t-T)}] \right)^\gamma,$$

substitute for J in (28), and divide out the common factor $F(t) \left(\frac{W}{1-\gamma} + \left[\frac{\eta}{\beta r} 1 - e^{r(t-T)} \right] \right)^\gamma$, then we derive a "reduced" equation for H ,

$$(99) \quad 0 = \frac{(1-\gamma)^2}{\gamma} \left(\frac{H}{\beta} \right) \frac{\gamma}{\gamma-1} + \frac{(1-\gamma)}{\gamma} \left(\frac{\dot{F}}{F} + H_t \right) + (1-\gamma)rH \\ + \frac{(1-\gamma)}{\gamma} \sum_1^m \alpha_i P_i H_i + \frac{(1-\gamma)}{2\gamma} \sum_1^m \sum_1^m \sigma_{ij} P_i P_j H_i H_j \\ + \sum_1^m (\alpha_i - r) P_i H_i + \frac{H}{2} \sum_1^m \sum_1^m v_{ij} (\alpha_i - r) (\alpha_j - r) \\ + \frac{1}{2H} \sum_1^m \sum_1^m \sigma_{ij} P_i P_j H_i H_j$$

and the associated optimal consumption and portfolio rules are

$$(100) \quad C^*(t) = \frac{(1-\gamma)}{\beta} \left[\left(\frac{H}{\beta} \right)^{\frac{1}{\gamma-1}} \left(\frac{W}{1-\gamma} + \frac{\eta}{\beta r} [1 - e^{r(t-T)}] \right) - \eta \right]$$

and

$$(101) \quad w_k^*(t)W = \left[\sum_1^m v_{jk} (\alpha_j - r) + \frac{H_k P_k}{H} \right] \left(\frac{W}{1-\gamma} + \frac{\eta}{\beta r} [1 - e^{r(t-T)}] \right),$$

$$k=1, \dots, m.$$

Although (99) is still a formidable equation from a computational point of view, it is less complex than the general equation (28), and it is possible to obtain an explicit solution for particular assumptions about the dependence of α_k and σ_k on the prices. Notice that both consumption and the asset demands are linear functions of wealth.

For a particular member of the HARA family, namely the Bernoulli logarithmic utility ($\gamma=0=\eta$ and $\beta=1-\gamma=1$) function, (28) can be solved in general. In this case, J will be of the form

$$(102) \quad J(W, P, t) = a(t) \log W + H(P, t) \text{ with } h(P, t) = a(T) = 0,$$

with $a(t)$ independent of the α_k and σ_k (and hence, the P_k). For the case when $F(t) \equiv 1$, we find $a(t) = T-t$ and the optimal rules become

$$(103) \quad C^* = \frac{W}{T-t}$$

and

$$(104) \quad w_k^* = \sum_{j=1}^m v_{kj} (\alpha_j - r) \quad , \quad k=1, \dots, m.$$

For the log case, the optimal rules are identical to those derived when α_k and σ_k were constants, with the understanding that the α_k and σ_k are evaluated at current prices. Hence, although we can solve this case for general price mechanisms, it is not an interesting one because different assumptions about price behavior have no effect on the decision rules.

The first of the alternative price mechanisms considered is called the "asymptotic 'normal' price-level" hypothesis which assumes that there exists a "normal" price function, $\bar{P}(t)$, such that

(105) $\lim_{t \rightarrow \infty} E_T [P(t)/\bar{P}(t)] = 1$, for $0 \leq T < t < \infty$,
i.e. independent of the current level of the asset price, the investor expects the "long-run" price to approach the normal price. A particular example which satisfies the hypothesis is that

$$(106) \quad \bar{P}(t) = \bar{P}(0)e^{vt}$$

and

$$(107) \quad \frac{dP}{P} = \beta \left[\phi + vt - \log(P(t)/P(0)) \right] dt + \sigma dz$$

where $\phi \equiv k + v/\beta + \sigma^2/4\beta$ and $k \equiv \log [\bar{P}(0)/P(0)]$.⁽³³⁾

For the purpose of analysis, it is more convenient to work with the variable $Y(t) \equiv \log [P(t)/P(0)]$ rather than $P(t)$. Substituting for P in (107) by using Itô's Lemma, we can write the dynamics for Y as

⁽³³⁾ In the notation used in previous sections, (107) corresponds to (5) with $\alpha(P,t) \equiv \beta[\phi + vt - \log(P(t)/P(0))]$. Note: "normal" does not mean "Gaussian" in the above use, but rather the normal long-run price of Alfred Marshall.

$$(108) \quad dY = \beta[\mu + \nu t - Y] dt + \sigma dz$$

where $\mu \equiv \phi - \sigma^2/2\beta$. Before examining the effects of this price mechanism on the optimal portfolio decisions, it is useful to investigate the price behavior implied by (106) and (107). (107) implies an exponentially-regressive price adjustment toward a normal price, adjusted for trend. By inspection of (108), Y is a normally-distributed random variable generated by a Markov process which is not stationary and does not have independent increments.⁽³⁴⁾ Therefore, from the definition of Y , $P(t)$ is log-normal and Markov. Using Itô's Lemma, one can solve (108) for $Y(t)$, conditional on knowing $Y(T)$, as

$$(109) \quad Y(t) - Y(T) = \left(k + \nu T - \frac{\sigma^2}{4\beta} - Y(T)\right) (1 - e^{-\beta\tau}) + \nu\tau + \sigma e^{-\beta t} \int_T^t e^{\beta s} dz$$

where $\tau \equiv t - T > 0$. The instantaneous conditional variance of $Y(t)$ is

$$(110) \quad \text{Var} [Y(t) | Y(T)] = \frac{\sigma^2}{2\beta} (1 - e^{-2\beta\tau}).$$

Given the characteristics of $Y(t)$, it is straightforward to derive the price behavior. For example, the conditional expected price can be derived from (110) and written as

$$(111) \quad E_T(P(t)/P(T)) = E_T \exp [Y(t) - Y(T)] \\ = \exp \left[\left(k + \nu T - \frac{\sigma^2}{4\beta} - Y(T)\right) (1 - e^{-\beta\tau}) + \nu\tau + \frac{\sigma^2}{4\beta} (1 - e^{-2\beta\tau}) \right]$$

It is easy to verify that (105) holds by applying the appropriate limit process to (111). Figure 2 illustrates the behavior of the conditional expectation mechanism over time.

⁽³⁴⁾ Processes such as (108) are called Ornstein-Uhlenbeck processes and are discussed, for example, in [3], p. 225.

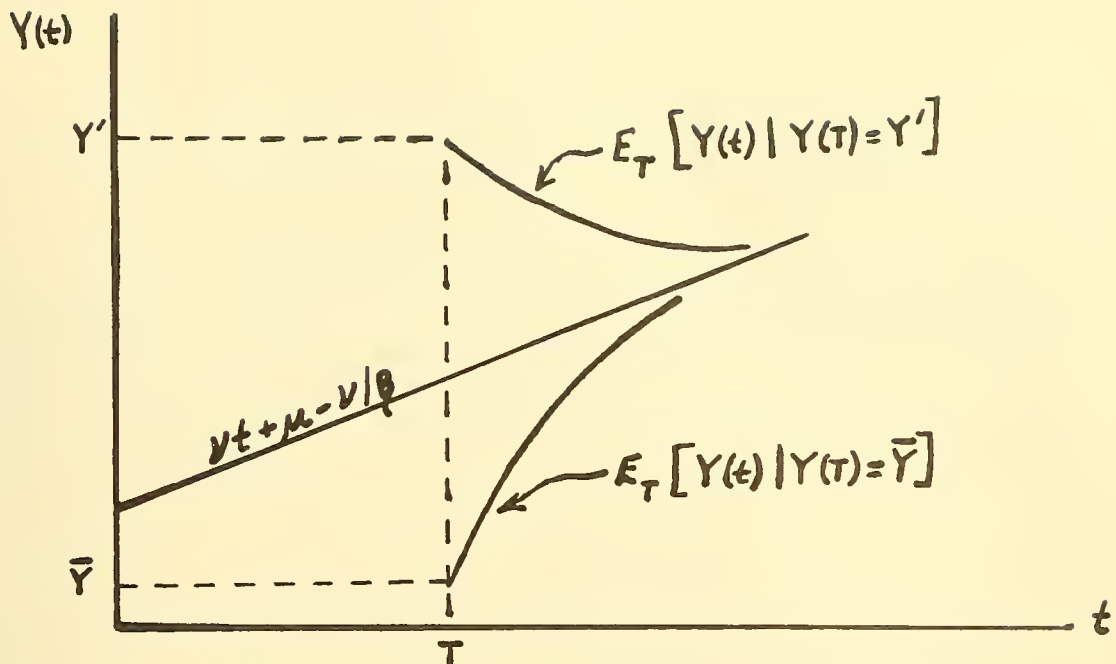


figure 2.

For computational simplicity in deriving the optimal consumption and portfolio rules, the two-asset model is used with the individual having an infinite time horizon and a constant absolute risk-aversion utility function, $U(C,t) = -e^{-C/\eta}$. The fundamental optimality equation then is written as

$$(112) \quad 0 = -e^{-\eta C^*}/\eta + J_t + J_W \left[w^* \left(\beta(\phi + vt - Y) - r \right) W + rW - C^* \right] \\ + \frac{1}{2} J_{WW} w^{*2} W^2 \sigma^2 + J_Y \beta(\mu + vt - Y) + \frac{1}{2} J_{YY} \sigma^2 + J_{YW} w^* W \sigma^2$$

and the associated equations for the optimal rules are

$$(113) \quad w^* W = -J_W [\beta(\phi + vt - Y) - r] / J_{WW} \sigma^2 - J_{YW} / J_{WW}$$

and

$$(114) \quad C^* = -\log(J_W)/\eta$$

Solving (112), (113), and (114), we write the optimal rules in explicit form as

$$(115) \quad w^* W = \frac{1}{\eta r \sigma^2} \left[\left(1 + \frac{\beta}{r} \right) (\alpha(P,t) - r) + \frac{\beta^2}{r^2} \left(\frac{\sigma^2}{2} + v - r \right) \right]$$

and

$$(116) \quad C^* = rW + \frac{\beta^2}{2\sigma^2\eta r} Y^2 - \frac{\beta}{\eta r \sigma^2} \left(\beta v t + \beta \phi - r + \beta \left(v + \frac{\sigma^2}{2} - r \right) \right) Y + a(t) \quad (35)$$

where $\alpha(P,t)$ is the instantaneous expected rate of return defined explicitly in footnote 33. To provide a basis for comparison, the solutions when the geometric Brownian motion hypothesis is assumed are presented as ⁽³⁶⁾

$$(117) \quad w^*W = \frac{(\alpha-r)}{\eta r \sigma^2}$$

and

$$(118) \quad C^* = rW + \frac{1}{\eta r} \left[\frac{(\alpha-r)^2}{2\sigma^2} - r \right]$$

To examine the effects of the alternative "normal price" hypothesis on the consumption-portfolio decisions, the (constant) α of (117) and (118) is chosen equal to $\alpha(P,t)$ of (115) and (116) so that, in both cases, the instantaneous expected return and variance are the same at the point of time of comparison. Comparing (115) with (117), we find that the proportion of wealth invested in the risky asset is always larger under the "normal price" hypothesis than under the geometric Brownian motion hypothesis. ⁽³⁷⁾ In particular, notice that even if $\alpha < r$, unlike in the geometric Brownian motion case, a positive amount of the risky asset is held. Figures 3a and 3b illustrate the behavior of the optimal portfolio holdings.

$$\begin{aligned} (35) \quad a(t) \equiv & \frac{1}{\eta} \left\{ \frac{r}{2\sigma^2} - 1 + \frac{\beta}{\sigma^2} \left(\phi - 1 - \frac{\sigma^2}{2r} \right) + \frac{\beta^2}{r\sigma^2} \left[\left(1 - \frac{\sigma^2}{2r} \right) \left(\phi + \frac{v}{r} + \frac{\sigma^2}{2r} - 1 \right) \right. \right. \\ & \left. \left. - \frac{\phi^2}{2} - \frac{\sigma^2}{2r} \right] + \frac{\beta v}{\sigma^2 r^2} \left(r + \beta - \beta \phi - \frac{\beta \sigma^2}{2r} \right) - \frac{\beta^2 v^2}{2\sigma^2 r^3} \right. \\ & \left. + \frac{\beta v t}{r \sigma^2} \left[r + \beta - \beta \phi - \frac{\beta \sigma^2}{2r} \right] - \frac{\beta^2 v^2 t^2}{2\sigma^2 r} \right\} . \end{aligned}$$

⁽³⁶⁾ For a derivation of (117) and (118), see [12], p.256.

⁽³⁷⁾ It is assumed that $v + \frac{\sigma^2}{2} > r$, i.e. the "long-run" rate of growth of the "normal" price is greater than the sure rate of interest so that something of the risky asset will be held in the short and long run.

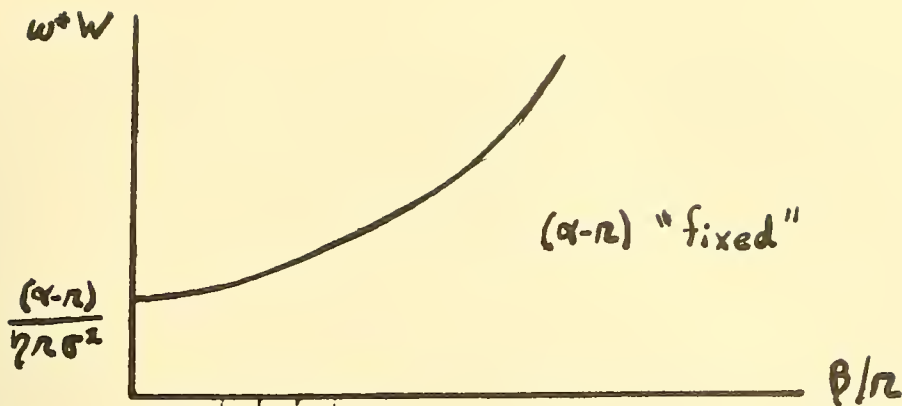


figure 3a.

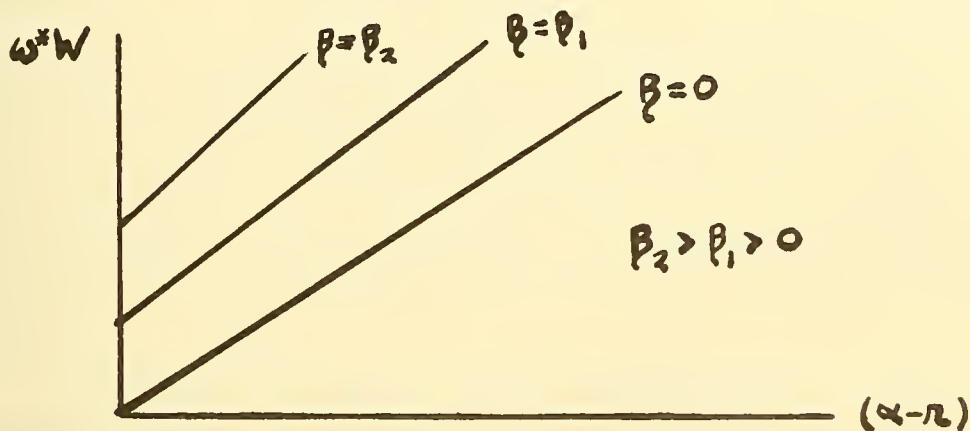


figure 3b.

The most striking feature of this analysis is that, despite the ability to make continuous portfolio adjustments, a person who believes that prices satisfy the "normal" price hypothesis will hold more of the risky asset than one who believes that prices satisfy the geometric Brownian motion hypothesis, even though they both have the same utility function and the same expectations about the instantaneous mean and variance.

The primary interest in examining these alternative price mechanism is to see the effects on portfolio behavior, and so, little will be said about the effects on consumption other than to present the optimal rule.

The second alternative price mechanism assumes the same type of price-dynamics equation as was assumed for the geometric Brownian motion: namely,

$$(119) \quad \frac{dP}{P} = \alpha dt + \sigma dz.$$

However, instead of the instantaneous expected rate of return, α , being a constant, it is assumed that α is itself generated by the stochastic differential equation

$$(120) \quad d\alpha = \beta(\mu - \alpha)dt + \delta \left(\frac{dP}{P} - \alpha dt \right) \\ = \beta(\mu - \alpha)dt + \delta \sigma dz.$$

The first term in (120) implies a long-run, regressive adjustment of the expected rate of return toward a "normal" rate of return, μ , where β is the speed of adjustment. The second term in (120) implies a short-run, extrapolative adjustment of the expected rate of return of the "error-learning" type, where δ is the speed of adjustment. I will call the assumption of a price mechanism described by (119) and (120) the "De Leeuw" hypothesis for Frank De Leeuw who first introduced this type mechanism to explain interest rate behavior.

To examine the price behavior implied by (119) and (120), we first derive the behavior of α , and then P . The equation for α , (120), is of the same type as (108) described previously. Hence, α is normally distributed and is generated by a Markov process. The solution of (120), conditional on knowing $\alpha(T)$ is

$$(121) \quad \alpha(t) - \alpha(T) = (\mu - \alpha(T)) (1 - e^{-\beta\tau}) + \delta \sigma e^{-\beta t} \int_T^t e^{\beta s} dz,$$

where $\tau \equiv t - T > 0$. From (121), the conditional mean and variance of $\alpha(t) - \alpha(T)$ are

$$(122) \quad E_T (\alpha(t) - \alpha(T)) = (\mu - \alpha(T)) (1 - e^{-\beta\tau})$$

and

$$(123) \quad \text{Var} [\alpha(t) - \alpha(T) | \alpha(T)] = \frac{\delta^2 \sigma^2}{2\beta} (1 - e^{-2\beta\tau}).$$

To derive the dynamics of P , note that, unlike α , P is not Markov although the joint process $[P, \alpha]$ is. Combining the results derived for $\alpha(t)$ with (119), we solve directly

for the price, conditional on knowing $P(T)$ and $\alpha(T)$,

$$(124) \quad Y(t) - Y(T) = \left(\mu - \frac{1}{2}\sigma^2 \right) \tau - \frac{(\mu - \alpha(T))}{\beta} (1 - e^{-\beta\tau}) \\ + \sigma \delta \int_T^t \int_T^s e^{-\beta(s-s')} dz(s') ds + \sigma \int_T^t dz,$$

where $Y(t) \equiv \log [P(t)]$. From (124), the conditional mean and variance of $Y(t) - Y(T)$ are

$$(125) \quad E_T [Y(t) - Y(T)] = \left(\mu - \frac{1}{2}\sigma^2 \right) \tau - \frac{(\mu - \alpha(T))}{\beta} (1 - e^{-\beta\tau})$$

and

$$(126) \quad \text{Var} [Y(t) - Y(T) | Y(T)] = \sigma^2 \tau + \frac{\sigma^2 \delta^2}{2\beta^3} \left[\beta\tau - 2(1 - e^{-\beta\tau}) \right. \\ \left. + \frac{1}{2}(1 - e^{-2\beta\tau}) \right] + \frac{2\delta\sigma^2}{\beta^2} \left[\beta\tau - (1 - e^{-\beta\tau}) \right].$$

Since $P(t)$ is log-normal, it is straightforward to derive the moments for $P(t)$ from (124), (125), and (126). Figure 4. illustrates the behavior of the expected price mechanism.

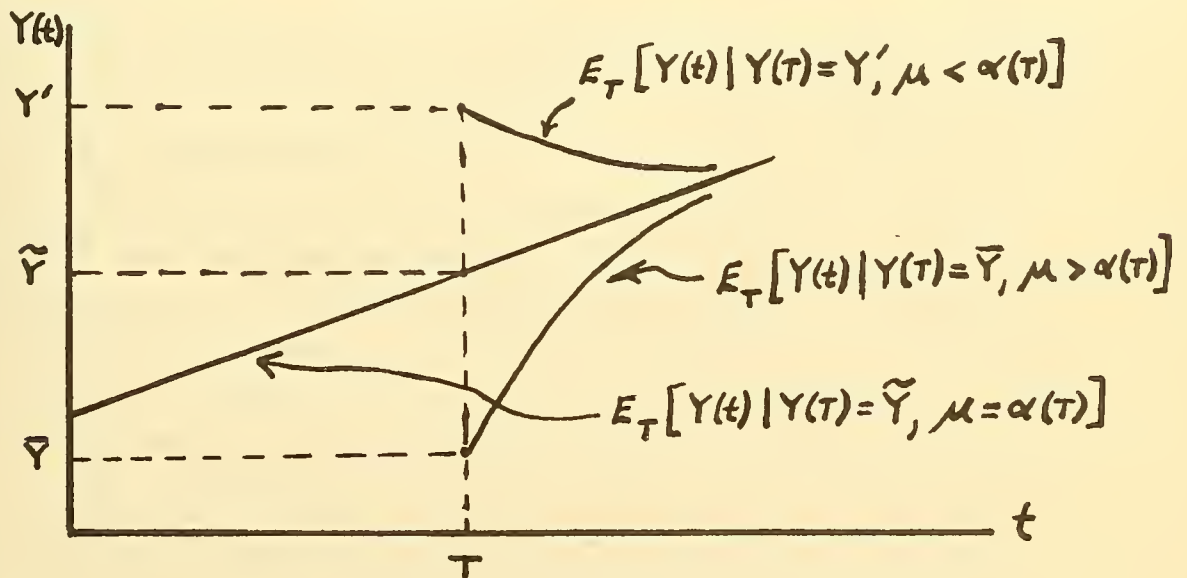


figure 4.

The equilibrium or "long-run" (i.e. $\tau \rightarrow \infty$) distribution for $\alpha(t)$ is stationary gaussian with mean μ and variance $\frac{\delta^2 \sigma^2}{2\beta}$, and the equilibrium distribution for $P(t)/P(T)$ is a stationary log-normal. Hence, the long-run behavior of prices under the De Leeuw hypothesis approaches the geometric Brownian motion.

Again, the two-asset model is used with the individual having an infinite time horizon and a constant absolute risk-aversion utility function, $U(C,t) = -e^{-\eta C}/\eta$. The fundamental optimality equation is written as

$$(127) \quad 0 = -\frac{e^{-\eta C^*}}{\eta} + J_t + J_W \left[w^*(\alpha-r)W + rW - C^* \right] \\ + \frac{1}{2} J_{WW} w^{*2} W^2 \sigma^2 + J_\alpha \beta(\mu-\alpha) + \frac{1}{2} J_{\alpha\alpha} \delta^2 \sigma^2 \\ + J_{W\alpha} \delta \sigma^2 w^* W$$

Notice that the state variables of the problem are W and α , which are both Markov, as is required for the dynamic programming technique. The optimal portfolio rule derived from (127) is,

$$(128) \quad w^* W = -\frac{J_W(\alpha-r)}{J_{WW} \sigma^2} - \frac{J_{W\alpha} \delta}{J_{WW}}$$

The optimal consumption rule is the same as in (114). Solving (127) and (128), the explicit solution for the portfolio rule is

$$(129) \quad w^* W = \frac{1}{\eta r \sigma^2 (r+2\delta+2\beta)} \left[(r + \delta + 2\beta)(\alpha-r) - \frac{\delta\beta(\mu-r)}{r+\delta+\beta} \right]$$

Comparing (129) with (117) and assuming that $\mu > r$, we find that under the De Leeuw hypothesis, the individual will hold a smaller amount of the risky asset than under the geometric Brownian motion hypothesis. Note also that $w^* W$ is a decreasing function of the long-run normal rate of return, μ . The interpretation of this result is that as μ increases for a given α , the probability increases that future " α 's" will be more favorable relative to the current α , and so there is a tendency to hold more of one's current wealth in the risk-free asset as a "reserve" for investment under more favorable conditions.

The last type of price mechanism examined differs from the previous two in that it is assumed that prices satisfy the geometric Brownian motion hypothesis. However, it is also

assumed that the investor does not know the true value of the parameter α , but must estimate it from past data. Suppose P is generated by equation (119) with α and σ constants, and the investor has price data back to time $-\tau$. Then, the best estimator for α , $\hat{\alpha}(t)$, is

$$(130) \quad \hat{\alpha}(t) = \frac{1}{t+\tau} \int_{-\tau}^t \frac{dP}{P}$$

where we assume, arbitrarily, that $\hat{\alpha}(-\tau) = 0$. From (130), we have that $E(\hat{\alpha}(t)) = \alpha$, and so, if we define the error term $\epsilon_t \equiv \alpha - \hat{\alpha}(t)$, then (119) can be re-written as

$$(131) \quad \frac{dP}{P} = \hat{\alpha}dt + \sigma d\hat{z}$$

where $d\hat{z} \equiv dz + \epsilon_t dt/\sigma$. Further, by differentiating (130) we have the dynamics for $\hat{\alpha}$, namely

$$(132) \quad d\hat{\alpha} = \frac{\sigma}{t+\tau} d\hat{z}.$$

Comparing (131) and (132) with (119) and (120), we see that this "learning" model is equivalent to the special case of the De Leeuw hypothesis of pure extrapolation (i.e. $\beta = 0$) where the degree of extrapolation (δ) is decreasing over time. If the two-asset model is assumed with an investor who lives to time T with a constant absolute risk-aversion utility function, and if (for computational simplicity) the risk-free asset is money (i.e. $r = 0$), then the optimal portfolio rule is

$$(133) \quad w^*W = \frac{(t+\tau)}{\eta\sigma^2} \log\left(\frac{T+\tau}{t+\tau}\right) \hat{\alpha}(t)$$

and the optimal consumption rule is

$$(134) \quad C^* = \frac{W}{T-t} - \frac{1}{\eta} \left[\log(T+\tau) + \frac{2}{T-t} \left(T-t - (T+\tau)\log(T+\tau) \right. \right. \\ \left. \left. + (t+\tau)\log(t+\tau) \right) + \frac{\hat{\alpha}^2}{2\sigma^2} \left(\frac{(t+\tau)^2}{(T-t)} \log\left[\frac{T+\tau}{t+\tau}\right] - \frac{(T-t)}{t+\tau} \right) \right]$$

By differentiating (133) with respect to t , we find that w^*W is an increasing function of time for $t < \bar{t}$, reaches a

maximum at $t = \bar{t}$, and then is a decreasing function of time for $\bar{t} < t < T$, where \bar{t} is defined by

(135) $\bar{t} = [T + (1-e)\tau]/e$. The reason for this behavior is that, early in life (i.e. for $t < \bar{t}$), the investor learns more about the price equation with each observation, and hence investment in the risky asset becomes more attractive. However, as he approaches the end of life (i.e. for $t > \bar{t}$), he is generally liquidating his portfolio to consume a larger fraction of his wealth, so that although investment in the risky asset is more favorable, the absolute dollar amount invested in the risky asset declines.

Consider the effect on (133) of increasing the number of available previous observations (i.e. increase τ). As expected, the dollar amount invested in the risky asset increases monotonically. Taking the limit of (133) as $\tau \rightarrow \infty$, we have that the optimal portfolio rule is

$$(136) \quad w^*W = \frac{(T-t)}{\eta\sigma^2} \alpha \quad \text{as } \tau \rightarrow \infty,$$

which is the optimal rule for the geometric Brownian motion case when α is known with certainty. Figure 5. illustrates graphically how the optimal rule changes with τ .

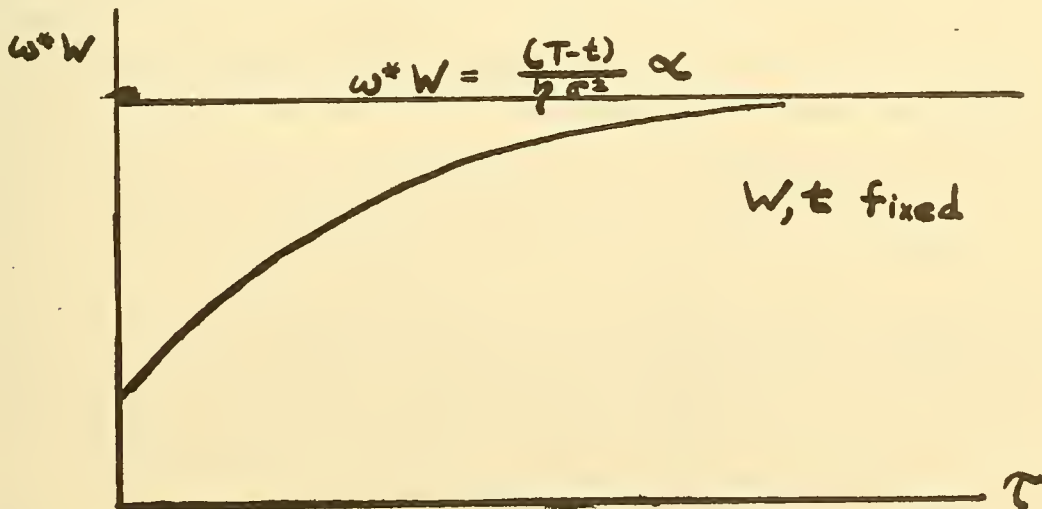


figure 5.

10. Conclusion. By the introduction of Itô's Lemma and the Fundamental Theorem of Stochastic Dynamic Programming (Theorem I), we have shown how to systematically construct and analyze optimal continuous-time dynamic models under uncertainty. The basic methods employed in studying the consumption-portfolio problem are applicable to a wide-class of economic models of decision making under uncertainty.

A major advantage of the continuous-time model over its discrete time analog is that one need only consider two types of stochastic processes: functions of Brownian motions and Poisson processes. This result limits the number of parameters in the problem and allows one to take full advantage of the enormous amount of literature written about these processes. Although I have not done so here, it is straightforward to show that the limits of the discrete-time model solutions as the period spacing goes to zero are the solutions of the continuous-time model.⁽³⁸⁾

A basic simplification gained by using the continuous-time model to analyze the consumption-portfolio problem is the justification of the Tobin-Markowitz portfolio efficiency conditions in the important case when asset price changes are stationarily and log-normally distributed. With earlier writers (Hakansson [6], Leland [10], Fischer [5], Samuelson [13], and Cass and Stiglitz [1]), we have shown that the assumption of the HARA utility function family simplifies the analysis and a number of strong theorems were proved about the optimal solutions. The introduction of stochastic wage income, risk of default, uncertainty about life expectancy, and alternative types of price dynamics serve to illustrate the power of the techniques as well as to provide insight into the effects of these complications on the optimal rules.

⁽³⁸⁾ For a general discussion of this result, see Samuelson [14].

References

- [1] Cass, D. and Stiglitz, J.E., "The Structure of Investor Preferences and Asset Returns, and Separability in Portfolio Allocation: A Contribution to the Pure Theory of Mutual Funds," Cowles Foundation Paper
- [2] Cootner, P., ed., The Random Character of Stock Market Prices, (Cambridge: The M.I.T. Press, 1964).
- [3] Cox, D.A. and Miller, H.D., The Theory of Stochastic Processes, (New York: John Wiley and Sons, Inc., 1968).
- [4] Dreyfus, S.E., Dynamic Programming and the Calculus of Variations, (New York: Academic Press, 1965).
- [5] Fischer, S., Essays on Assets and Contingent Commodities, Ph.D. Dissertation, Department of Economics, Massachusetts Institute of Technology, August, 1969.
- [6] Hakansson, N.H., "Optimal Investment and Consumption Strategies Under Risk for a Class of Utility Functions," forthcoming in Econometrica.
- [7] Itô, K., "On Stochastic Differential Equations," Mem. Amer. Math. Soc. No. 4 (1951).
- [8] _____ and McKean, H.P. Jr., Diffusion Processes and Their Sample Paths, (New York: Academic Press, 1964).
- [9] Kushner, H.J., Stochastic Stability and Control, (New York: Academic Press, 1967).
- [10] Leland, H.E., Dynamic Portfolio Theory, Ph.D. Dissertation Department of Economics, Harvard University, May, 1968.

- [11] McKean, H.P. Jr., Stochastic Integrals, (New York: Academic Press, 1969).
- [12] Merton, R.C., "Lifetime Portfolio Selection Under Uncertainty: The Continuous-time Case," Review of Economics and Statistics, LI, (August, 1969), p.247-257.
- [13] Samuelson, P.A., "Lifetime Portfolio Selection by Dynamic Stochastic Programming," Review of Economics and Statistics, LI, (August, 1969), p. 239-246.
- [14] _____, "The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances and Higher Moments," forthcoming in Review of Economic Studies.
- [15] Stratonovich, R.L., Conditional Markov Processes and Their Application to the Theory of Optimal Control, (New York: American Elsevier Publishing Co., 1968).

DEC 15 '75 DEC. 5 1987
AUG 11 '76
OCT 12 '76
JAN 11 '77
FEB 16 '77
MAY JUN 16 '77
MAY 26 '78
OCT 19 '78
OCT 13 '80
MAY 22 1985
OCT 6 1998
Lib-26-67

Lib-26-67

MIT LIBRARIES



3 9080 003 928 410

MIT LIBRARIES



3 9080 003 928 436

MIT LIBRARIES



3 9080 003 959 449

MIT LIBRARIES



3 9080 003 959 415

MIT LIBRARIES



3 9080 003 796 049

MIT LIBRARIES



3 9080 003 796 064

